Ecological Informatics 17 (2013) 73–81

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Modelling habitat preference, abundance and species richness of alien macrocrustaceans in surface waters in Flanders (Belgium) using decision trees

Pieter Boets *, Koen Lock, Peter L.M. Goethals

Laboratory of Environmental Toxicology and Aquatic Ecology, Ghent University, J. Plateaustraat 22, B-9000 Ghent, Belgium

ARTICLE INFO

Article history:
Received 15 April 2011
Received in revised form 2 March 2012
Accepted 2 June 2012
Available online 12 June 2012

Keywords:
Biological invasions
Classification trees
Habitat suitability modelling
Integrated modelling
Regression trees

ABSTRACT

The introduction and the spread of alien invasive species are a worldwide phenomenon causing global ecological and economic damages. Among the invaders, alien macrocrustaceans are known to be very successful invertebrates that colonise new habitats rapidly. Data from different fresh and brackish waters gathered by the Flemish Environment Agency (VMM) were used to build data-driven models predicting habitat preference, abundance and species richness of alien macro-Crustacea present in surface waters in Flanders. Different techniques such as regression and classification trees in combination with several optimisation methods (e.g. pruning) were used to construct the models. The performance of the models was moderate, because a balance between performance, ecological relevance and complexity was strived for. When using a three-fold cross validation it was found that the variation between the folds was limited, which is an indication of the robustness and the good reliability of the constructed models. Based on a sensitivity analysis the importance of conductivity, Kjeldahl nitrogen and shipping were stressed as well as graphically illustrated. Alien macrocrustaceans were predicted as present under brackish water conditions as well as in fresh waters with intensive ship traffic and low levels of organic pollution. The alien species richness was higher in rivers with intensive ship traffic and increased with increasing conductivity. Especially in brackish waters, alien macrocrustaceans reached high abundances. In fresh water, the abundance of alien species was generally lower. An integrated model that combined our habitat suitability model with a water quality model was used to predict the future distribution of alien macrocrustaceans. The predictions indicated that the prevalence and the species richness of alien macrocrustaceans are likely to increase with improving chemical water quality, whereas their abundance will probably decrease slightly. From our analysis, it is clear that models are a useful tool and that decision makers should focus on vulnerable areas such as brackish water areas and areas with intensive ship traffic in order to prevent the further introduction and spread of alien species.

1. Introduction

Rates of species' introductions are increasing globally as a consequence of increasing trade in the world (Vitousek et al., 1997). Unintentional intercontinental transport via ballast water and hull fouling of ships are key introduction vectors of aquatic species (Hulme et al., 2008). Changing environmental conditions, habitat degradation and the interconnection of waterways connecting previously separated biogeographic regions all promote the establishment and spread of alien species worldwide (Bij de Vaate et al., 2002; Boets et al., 2011a). The introduction of alien invasive species often has negative influences on native communities and ecosystems, with consequences such as species loss, biotic homogenization and changes in nutrient cycling (Gurevitch and Padilla, 2004; MacNeil et al., 2011). Therefore, techniques for modelling species' potential distributions could support pro-active strategies to avoid the introduction of alien species or to help in risk analysis by revealing those regions which are seen as hotspots for alien species introductions (Ba et al., 2010; Giovanelli et al., 2008; Worner and Gevrey, 2006). Besides their power to predict the future distribution of alien species, models can be used to assess the impact of alien invasive species on native species assemblages (Jaarsma et al., 2007). In this way, negative effects of alien invasive species on the environment (e.g. food web disturbance or habitat alteration) as well as on the economy (e.g. high costs for eradication and control) could be reduced and tackled in advance. Being able to determine which habitats are vulnerable for invasions is essential for a good management, because it is often either impossible or at least very expensive to eradicate alien invasive species after their establishment (Perrings et al., 2005).

Using conventional statistical multivariate methods to analyse data poses limits, because they are mainly applicable to linear data and have less flexibility in interpreting ecological data. Integrative and adaptive models that cover the non-linearity in a system are

* Corresponding author. Tel.: +32 472521819.
E-mail address: pieter.boets@ugent.be (P. Boets).

1574-9541/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecoinf.2012.06.001

envisioned in information processing in ecology (Park and Chon, 2007). Machine learning methods offer an advantage over traditional analysis techniques, because they do not introduce any prior assumptions about the relationship between variables (Džeroski and Drumm, 2003). Several data-driven modelling techniques such as Decision Trees, Artificial Neural Networks, Bayesian Belief Networks and Support Vector Machines have been proven to be successful in predicting the presence and distribution of species in aquatic ecosystems (Adriaenssens et al., 2004; Boets et al., 2010; Dominguez-Granda et al., 2011; Everaert et al., 2011; Goethals et al., 2007; Hoang et al., 2010). Selection of modelling techniques may be based on specific study objectives or the format of response variables (i.e. presence–absence versus abundance) and the availability and resolution of predictor variables such as climatic, physical–chemical and land-use data, which can be related to a species' occurrence and abundance. However, the selection among different modelling approaches is sometimes based on cost or convenience (Stohlgren et al., 2010). In this study, we opted to use classification and regression trees to predict the presence, abundance and species richness of alien macrocrustaceans in surface waters in Flanders, since these techniques are widely applied and yield results that are easy to interpret (Boets et al., 2010; Dakou et al., 2007; Everaert et al., 2011; Kampichler et al., 2010; Vaclavik and Meentemeyer, 2009).

In this paper, we focused on alien macrocrustaceans since these are widespread and represent, together with molluscs, the most important share of alien macroinvertebrates in many rivers across Europe (Bernauer and Jansen, 2006; Boets et al., 2011a; Messiaen et al., 2010; Nehring, 2006). Alien macrocrustaceans are often very successful in their new habitat. Their intrinsic characteristics, such as a short generation time, rapid growth with early sexual maturity, high fecundity and their euryhaline and omnivorous character make them extremely suitable for rapid expansion and establishment in freshwater ecosystems (Bij de Vaate et al., 2002). Several alien species belonging to the macrocrustaceans such as *Dikerogammarus villosus* or *Procambarus clarkii* are known to have an impact on native as well as alien biota (Boets et al., 2009, 2010). If we want to prevent the introduction and to reduce the impact of alien macrocrustaceans, a strict policy is needed. In this context, the models could help to support decision-making in water management by inducing measures for those regions which are at high risk.

In the present study, our goal was: (1) to predict in which habitats alien macrocrustaceans are likely to establish, (2) to determine which parameters positively or negatively influence the species richness of alien macrocrustaceans, (3) to assess which environmental conditions are favourable for alien macrocrustaceans to build up high densities and become dominant and (4) to make predictions on the future distribution of alien macrocrustaceans based on an integrated modelling approach. For the latter, habitat suitability models were combined with water quality models, which predict changes in chemical water quality due to the installation of planned wastewater treatment plants.

2. Materials and methods

2.1. Data collection

The dataset consisted of biological and chemical data collected by the Flemish Environment Agency (VVM). Since 1989, they monitor the water quality at more than 2,500 sampling locations scattered over different water bodies in Flanders (northern part of Belgium). In this way, a lot of data on macroinvertebrates as well as physical–chemical parameters was available. The model development was based on all samples collected during the year 2004, because this year was characterised by intensive sampling (over 800 samples) spread over Flanders. Macroinvertebrates were collected by standard handnet sampling or via artificial substrates if it was not possible to use the kick sampling method (Gabriels et al., 2010). Macroinvertebrates were identified to the level needed for the calculation of the Multimetric Macroinvertebrate Index Flanders (MMIF; Gabriels et al., 2010). Indigenous and alien species can belong to the same family and therefore, it was not clear from the VMM database if alien macrocrustaceans occurred in the samples. Since we wanted to make predictive models for alien macrocrustaceans, only species belonging to this group were identified to species level. In this way, data on the presence/absence, the abundance and the species richness of alien macrocrustaceans present per sampling location was available. Conductivity, pH and dissolved oxygen were always measured in the field during macroinvertebrate sampling. Other chemical parameters were retrieved from monitoring data. As the chemical monitoring, which was usually performed on a monthly basis, was not carried out simultaneously with the macroinvertebrate sampling, measurements from the last date before macroinvertebrate sampling were used. The slope of a watercourse was determined based on the difference in height between two points 1000 m apart, using GIS-software (version 9.3.1) applied on the Flemish Hydrographic Atlas (AGIV, 2006). The same data were used to determine the sinuosity on a stretch of 100 m. River morphology was evaluated based on pictures of the sampling sites: pool-riffle pattern and meandering were both quoted from 0 (absent) to 5 (well developed) and summed, which yielded a score from 0 to 10. Information on the number of passing ships on navigable waterways originated from the annual reports by nv De Scheepvaart and the River Information Services (RIS). For each sampling point, it was indicated whether ships passed or not and if so, how many ships passed on an annual base (based on the report of the year 2009). The complete dataset consisted of three response variables (presence/absence, number of alien macrocrustaceans and abundance of alien macrocrustaceans) and 16 predictor variables, two of which were discrete and 14 continuous (Table 1).

2.2. Model development and validation

Two types of decision trees were used to construct the models: classification and regression trees. A decision tree is called a classification tree (CT) if the response variable is qualitative (e.g. presence/absence of alien macrocrustaceans) and a regression tree (RT) if the response variable is quantitative (e.g. alien species richness or abundance). Decision trees were grown with a recursive partitioning algorithm from a training set of records, which is known as 'Top-Down Induction of Decision Trees' (Quinlan, 1986). For each step, the most informative input variable is selected as the root of the sub-tree and the current training set is split into subsets according to the values of the selected input variable. In this way, rules are generated that relate the predictor variables (e.g. river morphology) with the response variables (e.g. presence/absence of alien macrocrustaceans). For discrete predictor variables, a branch of a tree is typically created for each possible value of that particular variable. For continuous predictor variables, a threshold is selected and two branches are created based on that threshold. Tree construction ends when the variance of the class values of all examples in a node is within a certain range. Such nodes are called leaves and are labelled with a regression equation in case of regression trees or with the corresponding value of a class in case of classification trees (e.g. presence or absence).

Pruning was performed to prevent trees from over-fitting data (Džeroski and Drumm, 2003) and to make them easily interpretable (Dakou et al., 2007). Pruning can be used during tree construction (pre-pruning) and/or after the tree has been constructed (post-pruning). Pre-pruning is achieved when a minimum number of instances is needed before branching continues. Post-pruning on the other hand, implies that by changing the pruning confidence factor (PCF) some of the ending sub-trees of a highly branched tree can be replaced by leaves. In our case, pre- as well as post-pruning were performed as optimisation techniques.

Table 1
Average as well as minimum and maximum values (the range is indicated between brackets) of the assessed environmental parameters for the three constructed models: habitat preference, species richness and abundance of alien macrocrustaceans (BOD: Biological Oxygen Demand; COD: Chemical Oxygen Demand).

Variable	Unit	Habitat preference	Alien species richness	Abundance of alien species
Ammonium	mg N/L	1.82 (0.04–48.6)	1.38 (0.06–10.6)	1.49 (0.08–10.8)
BOD	mg/L	5.71 (1.0–354)	4.44 (1.06–30)	4.70 (1.0–30)
COD	mg/L	34.7 (2.5–680)	32.9 (8.0–132)	33.7 (2.5–132)
Dissolved oxygen	mg/L	7.0 (0.2–28.3)	7.07 (1.0–21.6)	7.11 (0.5–28.3)
Conductivity	µS/cm	1109 (90–17,570)	1612 (90–17,570)	1397 (90–17,570)
Kjeldahl nitrogen	mg N/L	3.9 (0.8–163)	2.66 (0.8–119)	2.89 (0.80–12.2)
Nitrate	mg N/L	3.3 (0.1–31.4)	2.7 (0.1–15.7)	3.1 (0.2–18)
Nitrite	mg N/L	0.16 (0.002–3.0)	0.14 (0.002–0.87)	0.15 (0.002–0.87)
Orthophosphate	mg P/L	0.46 (0.003–16.0)	0.38 (0.004–2.38)	0.38 (0.005–3.33)
Total phosphorus	mg P/L	1.06 (0.05–100)	0.82 (0.06–6.17)	0.82 (0.05–6.17)
pH		7.7 (6.0–9.4)	7.7 (6.4–9.2)	7.7 (6.0–9.4)
Sinuosity		1.04 (0.0–1.98)	1.03 (0.0–1.74)	1.05 (0.0–1.89)
Slope	m/1000 m	2.24 (0.0–42.5)	1.33 (0.0–20.5)	1.67 (0.0–20.5)
Number of ships		934 (0–32,772)	2032 (0–32,772)	1840 (0–32,772)
River morphology	classes (0–10)	3 (0–10)	3 (0–10)	3 (0–8)
Number of alien Crustacea	species/sample	0.44 (0–5)	1.2 (0–5)	1 (0–5)
Abundance of alien Crustacea	individuals/sample	10 (0–224)	25 (0–202)	22 (0–202)
Shipping	class (0,1)	Present (n = 107); Absent (n = 776)	Present (n = 30); Absent (n = 111)	Present (n = 51); Absent (n = 222)
Alien Crustacea	class (0,1)	Present (n = 299); Absent (n = 584)	Present (n = 94); Absent (n = 47)	Present (n = 182); Absent (n = 91)

The model training and evaluation was based on a three-fold cross validation. The dataset was, after reshuffling, randomly split in three subsets: two thirds were used for training and one third for validation. For each training and validation set a model was build and in this way, a performance value for each of the three different models was obtained. Average performance was used as final criterion for model evaluation. Model performance was based on the percentage Correctly Classified Instances (CCI) and Cohen's Kappa Statistic (K) for classification trees and the multiple correlation coefficient (R) for regression trees. In order to reach a satisfactory model performance, the CCI should be at least 70% and K should be at least 0.4 (Gabriels et al., 2007). For the multiple correlation coefficient, the closer the value is to 1, the better the model predicts the data (Everaert et al., 2010).

For the construction of the classification trees, the J48 algorithm was applied (Hall et al., 2009), which is a re-implementation of the C4.5 algorithm. Regression trees were built using M5′ (Wang and Witten, 1997), a re-implementation of the M5 algorithm (Quinlan, 1992). For both techniques, the standard settings from the machine learning package WEKA were applied (Witten and Frank, 2005), except for the PCF (pruning confidence factor) and the minimum number of instances required for further splits, which were adapted in order to obtain the most optimal model. The most optimal model was defined as a model with a good balance between a good technical performance (CCI, K) on the one hand and a high ecological relevance and reduced complexity on the other hand.

In total, 882 samples of the year 2004 from different sampling locations scattered over surface waters (different water types) in Flanders and comprising biological as well as physical–chemical and shipping data were used to build the models. All used response and predictor variables are listed in Table 1. Three different datasets were compiled, which were either used to predict habitat preference, abundance or species richness of alien macrocrustaceans. Classification trees, which were used to model habitat preference of alien macrocrustaceans can deal very well with missing data and outliers (Pham, 2006). Therefore, all data (882 instances) were used when applying this technique. When using regression trees, missing data and outliers (based on all values exceeding three times the standard deviation) were removed and the database was stratified, since this generally yields more consistent and robust performances (Everaert et al., 2010). Stratification implied that each possible outcome was represented by the same number of instances in the database. This resulted in total in 273 instances that were used for the abundance model and 141 instances that were used for the model predicting the species richness of alien macrocrustaceans.

2.3. Sensitivity analysis

Sensitivity analysis was done for the regression tree models to determine the weight of each variable in the regression equations as well as to check the robustness of the constructed models. For each parameter the minimum, maximum and average values were determined (Table 1). Afterwards, the outcome for each of these equations of the selected model was calculated by keeping all parameters constant (averages) except for the one which we wanted to analyse the sensitivity off, which ranged from its minimum to its maximum value. Dividing the maximum by the minimum outcome of the regression equation yielded a factor indicating the importance of each parameter. In addition, the effect of conductivity on the species richness and abundance of alien macrocrustaceans in relation to the other parameters was graphically illustrated for each of the three folds by keeping all other parameters constant (average values) except for conductivity which ranged from low (fresh water conditions) to high values (brackish water conditions).

2.4. Future dispersal

Predictions on the future prevalence, species richness and abundance of alien macrocrustaceans were made based on an integrated modelling approach. The constructed classification and regression tree models were combined with predictions on the improvement of the chemical water quality (PEGASE water quality model) due to the installation of planned wastewater treatment plants (Ronse and D'heygere, 2007). With the PEGASE model, physical–chemical data were modelled for 3 years: 2006 (reference data), 2015 and 2027, according to the deadlines set by the European Union Water Framework Directive (European Union, 2000). Based on previous research regarding the distribution of alien macrocrustaceans in Flanders (Boets et al., 2011a) and for practical reasons (e.g. due to the fact that not for all catchments in Flanders PEGASE data were available), we opted to investigate the distribution of alien macrocrustaceans in a selected catchment in Flanders (the canal Ghent–Terneuzen and its tributaries). Data generated per segment by the water quality model on physical–chemical water quality parameters were used as input for our habitat suitability models to make predictions on the future distribution. We assumed that shipping and number of ships remained constant, as we did not have predictive data on the future shipping intensity. Conductivity was kept constant as this parameter is not included in the PEGASE water quality model and no predictions on possible changes could be made. Finally, sinuosity and river morphology were kept constant as well, as these parameters are not expected to change in this timeframe. All other physical–chemical water quality parameters changed according to the water quality

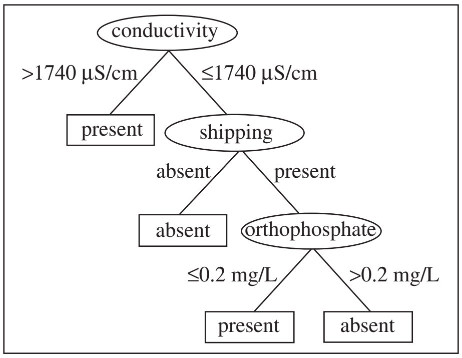


Fig. 1. Classification tree predicting the presence or absence of alien macrocrustaceans in surface waters in Flanders (pruning confidence factor = 0.25; Correctly Classified Instances = 72 ± 4%; Cohen's Kappa = 0.28 ± 0.06).

model. The final outcome for the different models was calculated and afterwards visualised in ArcMap (version 9.3.1).

3. Results

3.1. Habitat preference

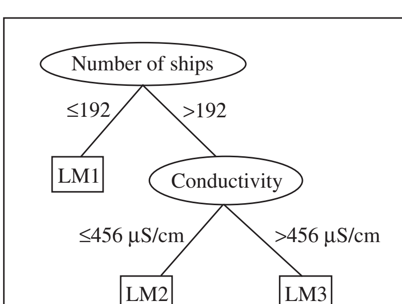
The presence or absence of alien macrocrustaceans could be accurately predicted based on the physical–chemical variables and the information regarding shipping. Only this tree that had an acceptable reliability in combination with a low complexity and a high ecological relevance was selected and presented (Fig. 1). After pruning (PCF = 0.25), a classification tree with four leaves was constructed. With this tree, 72 ± 4% of the instances were correctly classified and $K = 0.28 \pm 0.06$. The model revealed that alien macrocrustaceans are present at high conductivities (> 1740 µS/cm), which could be ascribed to brackish waters. If the conductivity was lower than or equalled 1740 µS/cm, other factors determined whether alien macrocrustaceans were present or not. In fresh water, alien macrocrustaceans were present in water with a low conductivity, where ships passed and where the orthophosphate concentration was lower than 0.2 mg/L (Fig. 1). This indicates that under freshwater conditions, shipping in combination with a good chemical water quality promotes the occurrence of alien macrocrustaceans.

3.2. Species richness

Similar to the model for habitat preference, only the model having an acceptable reliability in combination with a low complexity and a high ecological relevance is given in the results (Fig. 2). The model predicting the species richness of alien macrocrustaceans consisted of a regression tree with three leaves and an average performance of $R = 0.59 \pm 0.06$ (Fig. 2). The species richness of alien macrocrustaceans could be predicted as follows: if the number of ships was lower than or equal to 192 per year, the linear model 1 (LM1) was used. LM1 consisted of the variables chemical oxygen demand (COD), total phosphorus, conductivity, Kjeldahl nitrogen and slope. According to LM1, the number of alien macrocrustacean species present in the surface waters increased with increasing conductivity. If the number of ships was higher than 192 and the conductivity was lower than 456 µS/cm, the model LM2 was applied. LM2 used the same variables as LM1 and they contributed in a similar way to the increase or decrease of alien species richness. Finally, if the number of ships was higher than 192 and the conductivity was higher than 456 µS/cm, the alien species richness was determined by LM3. LM3 used the same variables as LM1 and LM2, but in this case, increasing conductivity as well as an increasing phosphorus concentration positively contributed to the established species richness of alien macrocrustaceans. Based on these regression equations, the average alien species richness was calculated for the different linear models. This resulted for LM1 (based on 112 sites) in 1.1 species, for LM2 (based on 5 sites) in 2.7 species and for LM3 (based on 21 sites) in 1.9 species. Sensitivity analysis indicated that when intensive ship traffic (> 192 ships per year) was present, the alien species richness was generally higher compared to low levels of ship traffic (< 192 ships per year) (Fig. 3A, B). Conductivity was an important variable determining the alien species richness, especially at higher conductivities, since after a certain threshold (456 µS/cm), the alien species richness increased with increasing conductivity (Fig. 3B). The highest species richness of alien macrocrustaceans was reached in freshwater with a good chemical water quality and intensive ship traffic. Sensitivity analysis of the linear regression equations pointed out that Kjeldahl nitrogen had a major contribution in determining the alien species richness. High levels of Kjeldahl nitrogen (> 10 mg/L) reduced the species richness of alien macrocrustaceans substantially.

3.3. Abundance

The most reliable and ecological relevant model yielded a regression tree with two leaves and a correlation coefficient (R) of 0.56 ± 0.08, which predicted the abundance of alien species in surface waters in Flanders (Fig. 4). If the conductivity was lower than or equal to 1041 µS/cm, the linear model LM1 was used. LM1 predicts the abundance of alien macrocrustaceans as a function of the variables ammonium, COD, conductivity, pH and the number of ships. Ammonium and conductivity have a negative influence on the abundance, whereas the abundance increases with increasing COD, the number of ships and pH. The average number of individuals calculated based on the 209 sites to which the equation could be applied was nine. In fresh waters, conductivity and ammonium negatively affect the abundance, which indicates that with increasing nutrient content, the abundance decreased. If the conductivity was higher than 1041 µS/cm, the linear model LM2 was used. LM2 used the variables ammonium, COD, conductivity, Kjeldahl nitrogen, orthophosphate, pH and number of ships to determine the abundance of alien macrocrustaceans. Only ammonium and Kjeldahl nitrogen negatively contributed to the abundance, whereas for all other predictor variables, the abundance increased when some



(LM1): Species richness = - 0.002 * COD - 0.295 * Total Phosphorus + 0.0001 * Conductivity - 0.106 * Kjeldahl Nitrogen - 0.080 * Slope + 1.660
(LM2): Species richness = - 0.005 * COD - 0.132 * Total Phosphorus + 0.0001 * Conductivity - 0.242 * Kjeldahl Nitrogen - 0.035 * Slope + 3.065
(LM3): Species richness = -0.005 * COD + 0.208 * Total Phosphorus + 0.0001 * Conductivity - 0.298 * Kjeldahl Nitrogen - 0.035 * Slope + 2.367

Fig. 2. Regression tree with regression equations predicting the species richness of alien macrocrustaceans in surface waters in Flanders (minimum number of instances = 4; correlation coefficient = 0.59 ± 0.06).

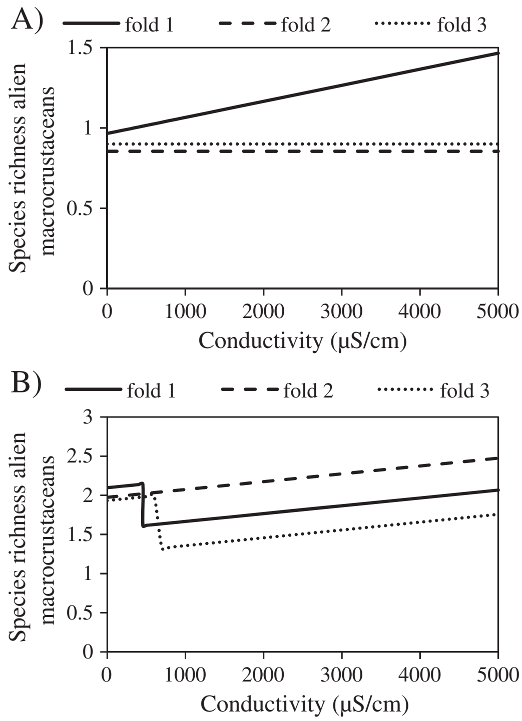


Fig. 3. Sensitivity analysis illustrating the effect of changing conductivity on the species richness of alien macrocrustaceans (A) at low (<192 ships per year) and (B) high levels of ship traffic (>192 ships per year) for all three folds.

or all of these variables increased. LM2 was applied to 64 sites with on average 64 individuals. This indicates that alien macrocrustaceans could reach higher abundances in brackish waters (high conductivity) and that high levels of orthophosphate do not necessarily negatively influence the abundance. In fresh water, increasing conductivity leads to a decrease in abundance of alien macrocrustaceans, whereas in brackish water increasing conductivity influenced the abundance positively (Fig. 5). Sensitivity analysis revealed that the regression equation of freshwater (LM1) was mostly influenced by the number of ships, whereas for brackish water (LM2) Kjeldahl nitrogen exerted an important negative effect on the abundance of alien macrocrustaceans.

3.4. Future dispersal

Based on our integrated model, we made predictions on the future prevalence, alien species richness and abundance of alien macrocrustaceans for the canal Ghent–Terneuzen and its tributaries. The model predicts an increase in prevalence of alien macrocrustaceans of 10% by the year 2027 (Fig. 6). The small canals around the city of Ghent as well as a tributary in the east are likely to be colonised by alien macrocrustaceans. There is also an increase in predicted alien species richness over the years (Fig. 7). Compared to the reference situation of 2006, the upper part of the canal Ghent–Terneuzen shows an increase in modelled alien species richness from on average one to on average two alien species by the year 2027. Although there is a general increase in the predicted prevalence and alien species richness, the abundance is predicted to decrease over the years. The predicted abundance in the canal Ghent–Terneuzen lies between 11 and 100 alien macrocrustacean individuals per sample in the year 2027 (Fig. 8). The newly colonised tributary in the east contains a very low predicted abundance of alien macrocrustaceans, with on average only one individual per sample.

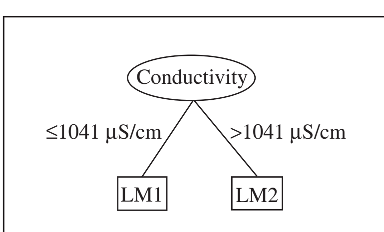
4. Discussion

4.1. Habitat modelling

Habitat suitability models can be applied to predict the potential distribution of alien species and to reveal their ecological niche preferences (Drake and Bossenbroek, 2009; Peterson, 2003; Pitt et al., 2009). These models can identify habitats at risk of invasion, which can help subsequent management efforts to maximize the efficacy of preventive measures to stop the spread of alien invasive species. Although these models are not 100% accurate in their predictions, they offer information regarding species preferences and their potential for invasion (Ba et al., 2010; Boets et al., 2010). Our developed habitat suitability models revealed that alien macrocrustaceans especially occur in brackish waters or freshwater with intensive ship traffic and low nutrient levels. Both salinity and shipping are known to be important parameters influencing the establishment and spread of alien macroinvertebrates (Ba et al., 2010; Bij de Vaate et al., 2002; Everaert et al., 2011; Grabowski et al., 2009).

Brackish waters are typically characterised by a low density and diversity of native species, hence it is easier for alien species to establish (Remane, 1958; Wolff, 1999). These waters with unsaturated ecological niches have a high potential to be invaded by alien macroinvertebrates (Paavola et al., 2005). Many alien macrocrustaceans are tolerant towards high salinities and therefore, they can easily establish in brackish water (Grabowski et al., 2007). In addition, it is assumed that brackish water species have a better chance of being transported alive than marine or freshwater species (Wolff, 1999). Grabowski et al. (2009) found that in the two largest rivers in the Baltic basin in Poland, alien amphipods were mostly found at conductivities above 800 µS/cm, where they reached high population densities. Moreover, brackish waters are subjected to a two sided 'invasion pressure', since species introduced in the freshwater as well as in the marine environment have the opportunity to migrate to these brackish waters (Nehring, 2006). Indigenous species inhabiting rivers have an increased conductivity risk to be displaced by alien ones, since alien species often have a competitive advantage over native species in polluted rivers (Grabowski et al., 2009). Organic discharges should therefore be minimized at all times, since these could ease the further spread of alien species (Grabowski et al., 2009).

Shipping is recognised as the most important vector of aquatic alien species introductions to Europe (Bij de Vaate et al., 2002; Gollasch, 2006). Improved ship design allows larger and faster ships, resulting in more frequent ship arrivals and larger amounts of ballast water being released. The construction of faster ships resulted in shorter



(LM1): Abundance = - 0.251 * Ammonium + 0.024 * COD - 0.013 * Conductivity + 8.412 * pH + 0.002 * Number of Ships - 49.994
(LM2): Abundance = - 0.666 * Ammonium + 0.697 * COD + 0.002 * Conductivity - 7.929 * Kjeldahl Nitrogen + 30.739 * Orthophosphate + 3.926 * pH + 0.0005 * Number of Ships + 0.4635

Fig. 4. Regression tree with regression equations predicting the abundance of alien macrocrustaceans in surface waters in Flanders (minimum number of instances = 4; correlation coefficient = 0.56 ± 0.08).

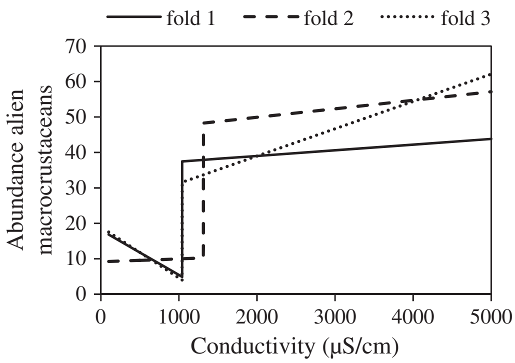


Fig. 5. Sensitivity analysis illustrating the effect of changing conductivity on the abundance of alien macrocrustaceans for all three folds.

voyages and consequently improved survival of alien species (Gollasch, 2006). Harbours and rivers with intensive ship traffic can be seen as hotspots for species introductions. Once an alien species is introduced into its new habitat, secondary dispersal via ballast water or hull fouling of ships can account for a substantial part of alien species dispersal (Bij de Vaate et al., 2002).

The model predicting the species richness of alien macrocrustaceans indicated that intensive ship traffic in combination with conductivity was the main factor determining a high alien species richness. As mentioned earlier, shipping is a key vector of species introductions (Bij de Vaate et al., 2002). Shipping activity is seen as an important proxy variable of propagule pressure (Ricciardi, 2006). The 'propagule pressure' concept focuses on the number of invading propagules for a given introduction and the frequency with which they are introduced (Williamson and Fitter, 1996). The more ships pass in a water body, the higher the chance that alien species become established and the higher the chance of an increased alien species richness. This phenomenon was also observed in the river Rhine, where the number of alien taxa decreased upstream with decreasing cargo transport (Wirth et al., 2010). Shipping is most intensive in human modified river ecosystems (e.g. large rivers and canals) with artificial or semi-natural embankments. These habitat conditions are often very attractive for alien species to establish and to become dominant. Boets et al. (2010) found that the alien invasive species *D. villosus* thrives very well in canals with artificial concrete riverbanks. Many alien macroinvertebrates can easily colonise and establish stable populations on these hard substrates (Van Riel et al., 2006) and are therefore preferred habitats.

Besides the possibility to reach a new habitat and become established, favourable environmental conditions are important to build up viable populations. The abundance of alien macrocrustaceans was mainly determined by conductivity. The average abundance calculated based on the linear regression equations was lower in freshwater (conductivity ≤ 1041 μS/cm) compared to brackish water (conductivity > 1041 μS/cm). It was found that in urban and densely populated areas, where high amounts of nutrients end up in river systems, the abundance of native species can be reduced and that of alien species increased (Vermonden et al., 2010). At their initial introduction stage, alien species can often have a competitive advantage at high nutrient concentrations (contributing to a low water quality) compared to indigenous species (Grabowski et al., 2007; Strayer, 2010). With increasing chemical as well as biological water quality, indigenous species might again be able to compete with alien species. Leuven et al. (2009) found that in urban waters in the Netherlands, indigenous macroinvertebrates were able to coexist and even dominate alien species in nutrient-poor, densely vegetated systems. However, our models indicate that also alien species, at least under freshwater conditions, can benefit from an improving water quality and that high nutrient concentrations can negatively affect the alien species richness. The high abundances detected at higher conductivities could be attributed to the presence of species like *Gammarus tigrinus*,

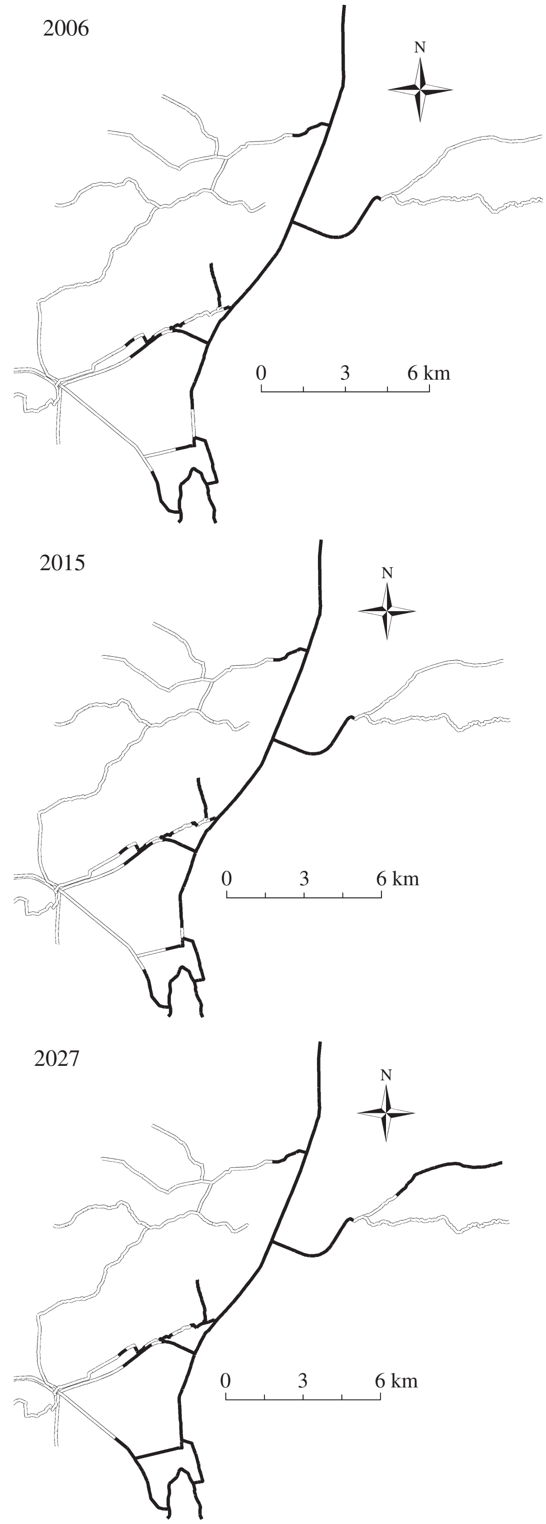


Fig. 6. Modelled prevalence of alien macrocrustaceans for the year 2006, 2015 and 2027, based on a water quality model for the canal Ghent–Terneuzen and its major tributaries (white = absent, black = present).

which especially occurs in brackish waters in Flanders, where it can reach high abundances (Boets et al., 2011b). This species shows a wide tolerance towards low or high levels of salinity, but is generally present in waters with a conductivity between 1200 and 3200 μS/cm. The high abundances of alien macrocrustaceans found at high conductivities could be ascribed to the high abundance of *G. tigrinus* present in these brackish waters.

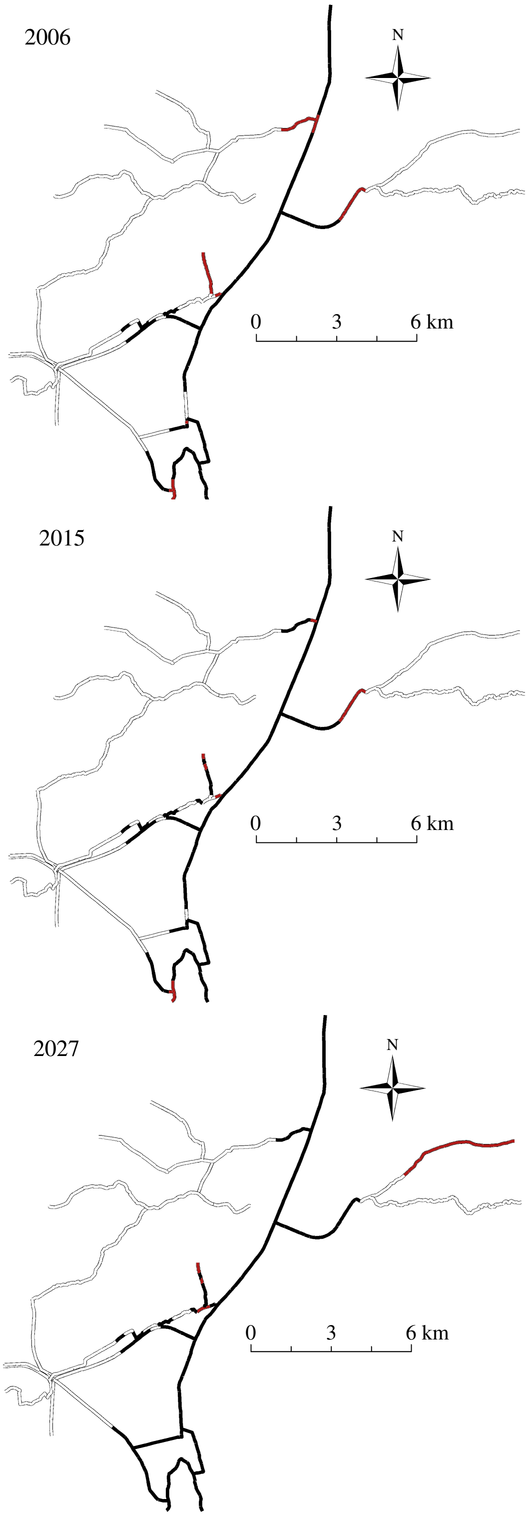


Fig. 7. Modelled species richness of alien macrocrustaceans for the year 2006, 2015 and 2027 based on a water quality model for the canal Ghent–Terneuzen and its major tributaries (white = no species, red = 1 species, black = 2 species).

4.2. Integrated modelling

Integrated models are often based on a combination of environmental and climatic conditions (e.g. Ficetola et al., 2007; Gallien et al., 2010). Habitats that meet these environmental and climatic constraints are identified as vulnerable for invasions. In this paper, a somewhat different approach was used: a habitat suitability model was integrated with a predictive water quality model. Based on such an integrated model, it is expected that by the year 2027, there will be an increase in prevalence and species richness of alien macrocrustaceans, but the abundance is predicted to decrease at some locations and to remain stable between 10 and 100 individuals per sample. At an initial stage, alien macrocrustaceans can have a competitive advantage in habitats with a low water quality and high nutrient levels. However, alien macrocrustaceans can also benefit from an improvement in the chemical water quality, especially in those watercourses that evolve from a bad to a moderate water quality. Due to the installation of wastewater treatment plants in Flanders, the chemical water quality is predicted to improve during the coming decades and therefore, our integrated model can give more accurate predictions on the future distribution of alien macrocrustaceans compared to simple habitat suitability models. Via this integrated model, valuable insight in the future potential invasive range of alien macrocrustaceans is given.

It has been suggested to incorporate species migration, population dynamics, biotic interactions and community ecology into species distribution models at multiple spatial scales (Guisan and Thuiller, 2005). In the next step, our integrated model could be optimized by including also dispersal of alien macrocrustaceans via a migration model. In our current model, species can only be present if an appropriate vector (shipping) is present, which imposes some limitations. The fact that species can actively colonise new areas and the time needed for this could be included in future models and give more accurate predictions. Recently, Gallardo et al. (2012) combined a large scale bioclimatic model with a local-scale migration model to predict the future distribution of *D. villosus* in Great Britain. Based on different scenarios of the annual migration speed, they were able to predict the dispersal of this alien macrocrustacean within the Great Ouse River catchment. They concluded that this approach helps to prevent and control the spread of alien invasive species and consequently can provide managers with a powerful spatial and temporal basis for informed decision-making.

4.3. Model performance

The performance of the habitat preference model was fair to moderate according to Gabriels et al. (2007). The variation on the different folds that were used was limited, which is an indication of the robustness of the constructed models. The fact that the performance was not very high could be due to some factors inherent to alien species. First of all, alien species may not yet have spread to all suitable habitats, making it difficult to determine species–environment relationships (Stohlgren et al., 2010). Secondly, alien species are often characterised as very opportunistic species, being able to easily cope with changes in environmental conditions (Nehring, 2006; Williamson and Fitter, 1996). Alien species can be seen as generalists, invading those niches which are available. Most alien species are omnivores and consequently, they do not pose any specific requirements regarding food availability (Strayer, 2010). All these elements make it difficult to accurately predict the habitat suitability and the distribution range of alien macrocrustaceans. Araújo and Guisan (2006) suggest that evaluation strategies should be discussed in the context of three possible uses: description, understanding and prediction. Complexity of model evaluation increases from explanation to prediction to the point where models that simply seek to describe a given pattern may not need to be evaluated, whereas the evaluation of models aiming at prediction is desirable but not always conceptually possible. Even though the accuracies of the models were not very high, Džeroski and Drumm (2003) state that, when using these techniques, we should bear in mind that the primary goal of such an analysis is to pinpoint the essential site characteristics rather than to predict the exact number of species. A major difficulty in applying data mining techniques for modelling alien invasive species is related to settings selection to obtain the most optimal model. The performance of models can be

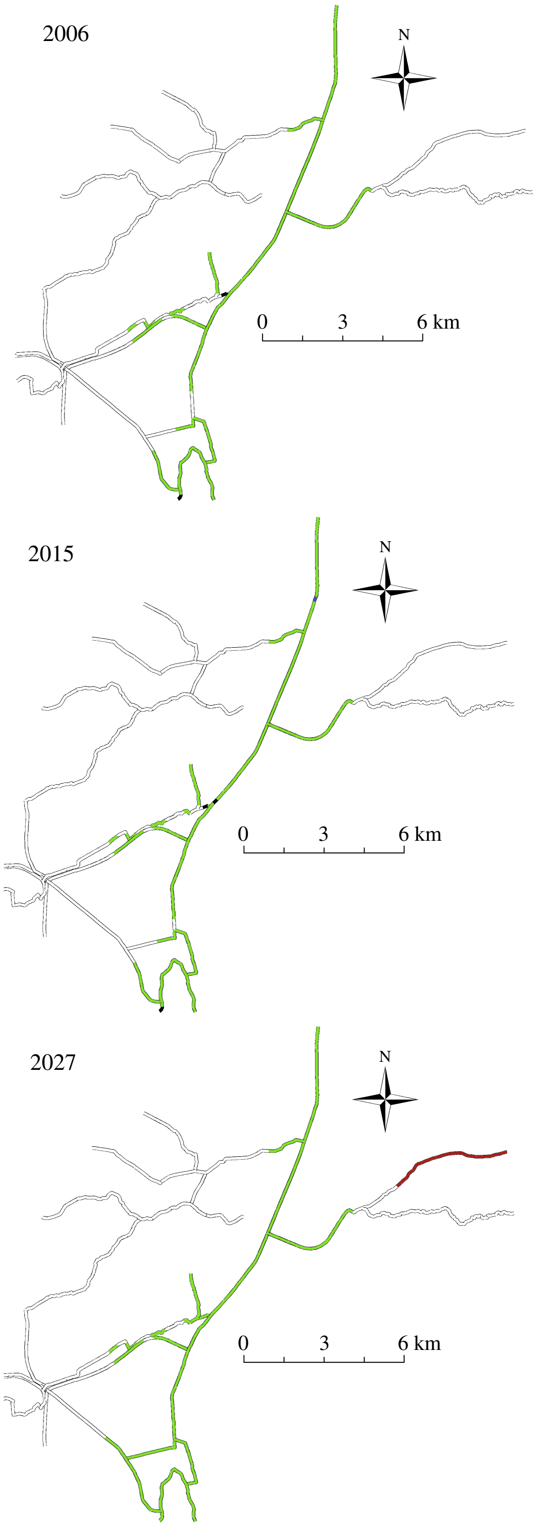


Fig. 8. Modelled abundance of alien macrocrustaceans for the year 2006, 2015 and 2027 based on a water quality model for the canal Ghent–Terneuzen and its major tributaries. Abundances were divided in several classes: white = absent, red = 1 individual, blue = 2–10 individuals, green = 11–100 individuals and black = 101–1000 individuals.

assessed from different perspectives, usually accounting for technical reliability, ecological relevance and user convenience. However, it is difficult to find a balance between these criteria, which are moreover to some extend both synergistic and antagonistic. Consequently, to compare and select optimal settings, there is a need for frameworks that can guide model developers in this selection, based on the specific characteristics of the data as well as the needs and interest of the model users (Willems, 2010). Ensemble forecasting models (Araújo and New, 2007; Stohlgren et al., 2010), using presence as well as absence data (Phillips et al., 2009), incorporating dispersal using estimates of dispersal rates (Midgley et al., 2006) or developing spatially-explicit species distribution models (Harris et al., 2009; Iverson et al., 2009; Smollik et al., 2010) have been suggested to overcome the shortcomings of traditional modelling techniques. Nevertheless, we can conclude that our models are useful and understandable for determining those environmental parameters and conditions that are important for alien macrocrustaceans to establish and become dominant. These models could be used by decision makers to pinpoint those regions within the aquatic environment that are under severe threat of invasion by alien species.

4.4. Conclusion

Alien macrocrustaceans have a preference for waters with a high conductivity. Shipping in combination with a good chemical water quality promotes the occurrence of alien macrocrustaceans. A maximum species richness of alien macrocrustaceans was reached in freshwater with a good chemical water quality and intensive ship traffic. In fresh water, increasing conductivity leads to a decrease in abundance of alien macrocrustaceans, whereas in brackish water increasing conductivity influenced the abundance positively. The predictions based on our integrated model approach indicated that the prevalence and the species richness of alien macrocrustaceans are likely to increase with improving chemical water quality, whereas their abundance will probably decrease slightly. From this study, it is clear that models are a useful tool for decision making and that policy makers should focus on vulnerable areas such as brackish water areas and areas with intensive ship traffic in order to prevent the further introduction and spread of alien species.

Acknowledgments

We would like to thank the Flemish Environment Agency (VMM) for the opportunity to study their samples and nv De Scheepvaart and the River Information System for providing us with data regarding shipping in Flanders. We also want to thank the VMM and Tom D'heygere in particular for providing the PEGASE water quality data. Koen Lock was supported by a post-doctoral fellowship from the Fund for Scientific Research (FWO-Vlaanderen, Belgium).

References

Adriaenssens, V., Goethals, P.L.M., Charles, J., De Pauw, N., 2004. Application of Bayesian Belief Networks for the prediction of macroinvertebrate taxa in rivers. Annales de Limnologie — International Journal of Limnology 40, 181–191.

AGIV, Agentschap voor Geografische Informatie Vlaanderen, 2006. Vlaamse Hydrografische Atlas. http://geo-vlaanderen.agiv.be/geo-vlaanderen/vha2006.

Araújo, M.B., Guisan, A., 2006. Five (or so) challenges for species distribution modeling. Journal of Biogeography 33, 1677–1688.

Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends in Ecology & Evolution 22, 42–47.

Ba, J.W., Hou, Z.E., Platvoet, D., Zhu, L., Li, S., 2010. Is *Gammarus tigrinus* (Crustacea, Amphipoda) becoming cosmopolitan through shipping? Predicting its potential invasive range using ecological niche modeling. Hydrobiologia 649, 183–194.

Bernauer, D., Jansen, W., 2006. Recent invasions of alien macroinvertebrates and loss of native species in the upper Rhine River, Germany. Journal of Aquatic Invasions 1, 55–71.

Bij de Vaate, A., Jazdzewski, K., Ketelaars, H.A.M., Gollasch, S., Van der Velde, G., 2002. Geographical patterns in range extension of Ponto-Caspian macroinvertebrate species in Europe. Canadian Journal of Fisheries and Aquatic Sciences 59, 1159–1174.

Boets, P., Lock, K., Cammaerts, R., Plu, D., Goethals, P.L.M., 2009. Occurrence of the invasive crayfish *Procambarus clarkii* (Girard, 1852) in Belgium (Crustacea: Cambaridae). Belgian Journal of Zoology 139, 173–176.

Boets, P., Lock, K., Messiaen, M., Goethals, P.L.M., 2010. Combining datadriven methods and lab studies to analyse the ecology of *Dikerogammarus villosus*. Ecological Informatics 5, 133–139.

Boets, P., Lock, K., Goethals, P.L.M., 2011a. Using long-term monitoring to investigate the changes in species composition in the harbour of Ghent (Belgium). Hydrobiologia 663, 155–166.
Boets, P., Lock, K., Goethals, P.L.M., 2011b. Alterations in the gammarid (Amphipoda) fauna in brackish polder waters of Flanders (Belgium). Journal of Crustacean Biology 31, 270–277.
Dakou, E., D'heygere, T., Dedecker, A.P., Goethals, P.L.M., Dimitriadou, M.L., De Pauw, N., 2007. Decision tree models for prediction of macroinvertebrate taxa in the river Axios (Northern Greece). Aquatic Ecology 41, 399–411.
Dominguez-Granda, L., Lock, K., Goethals, P.L.M., 2011. Application of classification trees to determine biological and chemical indicators for river assessment: case-study in the Chaguana watershed (Ecuador). Journal of Hydroinformatics 13, 489–499.
Drake, J.M., Bossenbroek, J.M., 2009. Profiling ecosystem vulnerability to invasion by zebra mussels with support vector machines. Theoretical Ecology 2, 189–198.
Džeroski, S., Drumm, D., 2003. Using regression trees to identify the habitat preference of the sea cucumber (*Holothuria leucospilota*) on Rarotonga, Cook Islands. Ecological Modelling 170, 219–226.
European Union, 2000. European Union Water Framework Directive. Directive 2000/60/EC of the European Parliament.
Everaert, G., Pauwels, I., Goethals, P.L.M., 2010. Development of data-driven models for the assessment of macroinvertebrates in rivers in Flanders. Proceedings of the International Environmental Modelling and Software Society (iEMSs) Fifth Biennial Meeting: International Congress on Environmental Modelling and Software, Ottawa, Canada.
Everaert, G., Boets, P., Lock, K., Goethals, P.L.M., 2011. Application of decision trees to analyze the ecological impact of invasive species in polder lakes in Flanders, Belgium. Ecological Modelling 222, 2202–2212.
Ficetola, G.F., Thuiller, W., Miaud, C., 2007. Prediction and validation of the potential global distribution of a problematic alien invasive species — the American bullfrog. Diversity and Distributions 13, 476–485.
Gabriels, W., Goethals, P.L.M., Dedecker, A., Lek, S., De Pauw, N., 2007. Analysis of macrobenthic communities in Flanders, Belgium, using a stepwise input variable selection procedure with artificial neural networks. Aquatic Ecology 41, 427–441.
Gabriels, W., Lock, K., De Pauw, N., Goethals, P.L.M., 2010. Multimetric Macroinvertebrate Index Flanders (MMIF) for biological assessment of rivers and lakes in Flanders (Belgium). Limnologica 40, 199–207.
Gallardo, B., Errea Paz, M., Aldridge, D.C., 2012. Application of bioclimatic models coupled with network analysis for risk assessment of the killer shrimp, *Dikerogammarus villosus*, in Great Britain. Biological Invasions, http://dx.doi.org/10.1007/s10530-011-0154-0.
Gallien, L., Munkemuller, T., Albert, C.H., Boulangeat, I., Thuiller, W., 2010. Predicting potential distributions of invasive species: where to go from here? Diversity and Distributions 16, 331–342.
Giovanelli, J.G.R., Haddad, C.F.B., Alexandrino, J., 2008. Predicting the potential distribution of the alien invasive American bullfrog (*Lithobates catesbeianus*) in Brazil. Biological Invasions 10, 585–590.
Goethals, P.L.M., Dedecker, A.P., Gabriels, W., Lek, S., De Pauw, N., 2007. Applications of artificial neural networks predicting macroinvertebrates in freshwaters. Aquatic Ecology 41, 491–508.
Gollasch, S., 2006. Overview on introduced aquatic species in European navigational and adjacent waters. Helgoland Marine Research 60, 84–89.
Grabowski, M., Bacela, K., Konopacka, A., 2007. How to be an invasive gammarid (Amphipoda: Gammaroidea) — comparison of life history traits. Hydrobiologia 590, 75–84.
Grabowski, M., Bacela, K., Konopacka, A., Jazdzewski, K., 2009. Salinity related distribution of alien amphipods in rivers provides refugia for native species. Biological Invasions 11, 2107–2117.
Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8, 993–1009.
Gurevitch, J., Padilla, D.K., 2004. Are invasive species a major cause of extinctions? Trends in Ecology & Evolution 19, 470–474.
Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The WEKA data mining software: an update. SIGKDD Explorations 11.
Harris, C.M., Park, K.J., Atkinson, R., Edwards, C., Travis, J.M.J., 2009. Invasive species control: incorporating demographic data and seed dispersal into a management model for *Rhododendron ponticum*. Ecological Informatics 4, 226–233.
Hoang, T.H., Lock, K., Mouton, A., Goethals, P.L.M., 2010. Application of decision trees and support vector machines to model the presence of macroinvertebrates in rivers in Vietnam. Ecological Informatics 5, 140–146.
Hulme, P.E., Bacher, S., Kenis, M., Klotz, S., Kuehn, I., Minchin, D., Nentwig, W., Olenin, S., Panov, V., Pergl, J., Pysek, P., Roques, A., Sol, D., Solarz, W., Vila, M., 2008. Grasping at the routes of biological invasions: a framework for integrating pathways into policy. Journal of Applied Ecology 45, 403–414.
Iverson, L.R., Prasad, A.M., Bossenbroek, J., Sydnor, D., Schwartz, M.W., 2009. Modeling potential movements of an ash threat: the emerald ash borer. In: Pye, J., Raucher, M. (Eds.), Advances in Threat Assessment and Their Application to Forest and Rangeland Management (http://www.threats.forestencyclopedia.net).
Jaarsma, N.G., Bergman, M., Schulze, F.H., Bij de Vaate, A., 2007. Macro-invertebrates in a dynamic river environment: analysis of time series from artificial substrates, using a 'white box' neural network modelling method. Aquatic Ecology 41, 413–425.
Kampichler, C., Wieland, R., Calmé, C., Weissenberger, H., Arriaga-Weiss, S., 2010. Classification in conservation biology: a comparison of five machine-learning methods. Ecological Informatics 5, 441–450.
Leuven, R.S.E.W., van der Velde, G., Baijens, I., Snijders, J., van der Zwart, C., Lenders, R., Bij de Vaate, A., 2009. The river Rhine: a global highway for dispersal of aquatic invasive species. Biological Invasions 11, 1989–2008.
MacNeil, C., Platvoet, D., Briffa, M., 2011. Direct and indirect effects of species displacements: an invading freshwater amphipod can disrupt leaf-litter processing and shredder efficiency. Journal of the North American Benthological Society 30, http://dx.doi.org/10.1899/10-056.1.
Messiaen, M., Lock, K., Gabriels, W., Vercauteren, T., Wouters, K., Boets, P., Goethals, P.L.M., 2010. Alien macrocrustaceans in freshwater ecosystems in the eastern part of Flanders (Belgium). Belgian Journal of Zoology 140, 30–39.
Midgley, G.F., Hughes, G.O., Thuiller, W., Rebelo, A.G., 2006. Migration rate limitations on climate change-induced range shifts in Cape Proteaceae. Diversity and Distributions 12, 555–562.
Nehring, S., 2006. Four arguments why so many alien species settle into estuaries, with special reference to the German River Elbe. Helgoland Marine Research 60, 127–134.
Paavola, M., Olenin, S., Leppakoski, E., 2005. Are invasive species most successful in habitats of low native species richness across European brackish water seas? Estuarine, Coastal and Shelf Science 64, 738–750.
Park, Y.S., Chon, T.S., 2007. Biologically-inspired machine learning implemented to ecological informatics. Ecological Modelling 203, 1–7.
Perrings, C., Dehnen-Schmutz, K., Touza, J., Williamson, M., 2005. How to manage biological invasions under globalization. Trends in Ecology & Evolution 20, 212–215.
Peterson, A.T., 2003. Predicting the geography of species' invasions via ecological niche modeling. The Quarterly Review of Biology 78, 419–433.
Pham, H., 2006. Springer Handbook of Engineering Statistics. Springer, London, pp. 551–569 (Chapter 30).
Phillips, S.J., Dudik, M., Elith, J., Graham, C., Lehmann, A., Leathwick, J., Ferrier, S., 2009. Sample selection bias and presence-only models of species distributions. Ecological Applications 19, 181–197.
Pitt, J.P.W., Worner, S.P., Suarez, A.V., 2009. Predicting Argentine ant spread over the heterogeneous landscape using a spatially explicit stochastic model. Ecological Applications 19, 1176–1186.
Quinlan, J.R., 1986. Induction of decision trees. Machine Learning 1, 81–106.
Quinlan, J.R., 1992. Learning with continuous classes. Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence. World Scientific Singapore, pp. 343–348.
Remane, A., 1958. Oekologie des Brackwassers. Die Binnengewaesser, Stuttgart. (22 pp.).
Ricciardi, A., 2006. Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. Diversity and Distributions 12, 425–433.
Ronse, Y., D'heygere, T., 2007. Waterkwaliteitsmodellering als beleidsondersteunend instrument bij de opmaak van stroomgebiedbeheerplannen. Water 32 (6 pp.).
Smollik, M.G., Dullinger, S., Essl, F., Kleinbauer, I., Leitner, M., Peterseil, J., Stadler, L.-M., Vogl, G., 2010. Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. Journal of Biogeography 37, 411–422.
Stohlgren, T.J., Ma, P., Kumar, S., Rocca, M., Morisette, J.T., Jarnevich, C.S., Benson, N., 2010. Ensemble habitat mapping of invasive plant species. Risk Analysis 30, 224–235.
Strayer, D.L., 2010. Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. Freshwater Biology 55, 152–174.
Vaclavik, T., Meentemeyer, R.K., 2009. Invasive species distribution modeling (iSDM): are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling 220, 3248–3258.
van Riel, M.C., van der Velde, G., Bij de Vaate, A., 2006. To conquer and persist: colonization and population development of the Ponto-Caspian amphipods *Dikerogammarus villosus* and *Chelicorophium curvispinum* on bare stone substrate in the main channel of the River Rhine. Archives of Hydrobiology 166, 23–39.
Vermonden, K., Leuven, R.S.E.W., van der Velde, G., 2010. Environmental factors determining invasibility of urban waters for exotic macroinvertebrates. Diversity and Distributions 16, 1009–1021.
Vitousek, P.M., Dantonio, C.M., Loope, L.L., Réjmanek, M., Westbrooks, R., 1997. Introduced species: a significant component of human-caused global change. New Zealand Journal of Ecology 21, 1–16.
Wang, Y., Witten, I.H., 1997. Induction of model trees for predicting continuous classes. Proceedings of the Poster Papers of the European Conference on Machine Learning. Faculty of Informatics and Statistics, University of Economics, Prague.
Willems, W., 2010. Habitat suitability models for the analysis and prediction of macrobenthos of the North Sea. PhD thesis, Ghent University. 274 pp.
Williamson, M.H., Fitter, A., 1996. The characters of successful invaders. Biological Conservation 78, 163–170.
Wirth, A., Schmera, D., Baur, B., 2010. Native and alien macroinvertebrate richness in a remnant of the former river Rhine: a source for recolonisation of restored habitats? Hydrobiologia 652, 89–100.
Witten, I.H., Frank, E., 2005. Data Mining: Practical Machine Learning Tools and Techniques, second edition. Morgan Kaufmann Publishers, San Francisco. (525 pp.).
Wolff, W.J., 1999. Exotic invaders of the meso-oligohaline zone of estuaries in the Netherlands: why are there so many? Helgolander Meeresunters 52, 393–400.
Worner, S.P., Gevrey, M., 2006. Modelling global insect species assemblages to determine risk of invasion. Journal of Applied Ecology 43, 858–867.